

NOAA's Integrated Observations and Data Management and Strategic Portfolio Tools

Martin Yapur & Eric Miller
NOAA/NESDIS
1335 East-West Highway, SSMC-1
Silver Spring, MD 20910 USA

Abstract - The National Oceanic and Atmospheric Administration is developing tools to support the analysis and acquisition of architectures for integrated observations and data management strategic portfolios. One of these tools is the "CasaNOSA Analysis Tool" (CASRT), a desktop software tool that retrieves and matches observing requirements and capabilities, and uses a variety of algorithms to measure the degree of fit or performance between systems and requirements at the attribute level. This tool pulls data from the database of NOAA observing and data management requirements and capabilities that NESDIS' Office of Systems Development has compiled over the last five years. NOAA's operational and science-oriented programs have more than 800 mission-critical observing requirements. NOAA operates over 80 different land, sea, and space-based observing systems, and also obtains data from a wide range of public and private sector sources. NOAA's observing requirements and capabilities are defined in terms of key attributes such as geographic coverage, horizontal or vertical resolution, measurement accuracy, and re-visit frequency. The NOAA database of requirements and capabilities resides in an open source, web-based repository system and is managed through a suite of collaborative tools. Results from these CASRT gap analyses will be used for investment portfolio analysis, decision support, statistical analysis, and enterprise architecture modeling.

I. INTRODUCTION

For over three decades, Earth observations have been fundamental to the National Oceanic and Atmospheric Administration (NOAA's) mission. NOAA relies on its observing systems for virtually every activity the agency does, from research to long range forecasting, to day-to-day regulatory decisions. [1] Critical to NOAA's mission are the environmental data from satellites, ships, aircraft, buoys, and other remote and in-situ observing systems. Exploiting these vital systems requires specialized communication, data management, and high-performance computing systems to process, code, transmit, receive, model, share, and store data. These capabilities enable NOAA to distribute the information in the right format, to the right users, at the right time as well as archive it for future use. NOAA is focused to build an operational earth observation system that is comprehensive, coordinated, and sustained. [2]

II. GLOBAL AWARENESS

Many nations are aware of the increasing social and economic need for an effective global Earth observation system to realize environmental and scientific benefits. As of July 2008, 75 nations and the European Commission have joined the Group on Earth Observations (GEO). NOAA is working with other national partners and GEO to strengthen world wide cooperation for Earth observations that address three imperatives (social, economical, and scientific) across four spatial domains (space, oceans, atmosphere, and land). [2] Human knowledge of the Earth's systems, although advanced in numerous areas, is far from complete. Current efforts to observe and understand the Earth environment must progress from observation systems of today to coordinated, timely, accurate, sustained global information-developed in accordance with compatible standards that lead to sound decisions and actions. Today, global Earth observation data are limited by accessibility, eroding technical infrastructure, spatial and temporal gaps, inadequate data integration and interoperability, continuity of observations, inadequate user involvement, lack of data process systems, and long term data archiving.

III. NOAA's STRATEGIC PORTFOLIO

In order for NOAA to maintain a high degree of data gathering and informational excellence, and best address increasing demands for environmental information, NOAA needs to construct and maintain a better integrated, cost-efficient architecture of observing and data systems. This end-to-end architecture will more effectively satisfy user needs, through the definition and prioritization of observing requirements, and investments in the most cost-effective alternatives for observing systems, product generation; and archival and access systems. A better integrated data management architecture will more effectively support the NOAA's mission objectives with common and NOAA approved hardware and software, policies, practices, and standards. In order to meet today's ever-growing public needs, we must conduct our business using more efficient methods. NOAA's goal in this pursuit is to have an integrated, comprehensive and sustained observation and data management system serving NOAA's four mission goals (Weather and Water, Commerce and Transportation, Climate, and Ecosystems) by improving our ability to

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE SEP 2008		2. REPORT TYPE		3. DATES COVERED 00-00-2008 to 00-00-2008	
4. TITLE AND SUBTITLE NOAA's Integrated Observations and Data Management and Strategic Portfolio Tools				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Oceanic and Atmospheric Administration, NOAA/NESDIS, 1335 East-West Highway, SSMC-1, Silver Spring, MD, 20910				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES See also ADM002176. Presented at the MTS/IEEE Oceans 2008 Conference and Exhibition held in Quebec City, Canada on 15-18 September 2008.					
14. ABSTRACT see report					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 6	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

understand, describe, and predict changes to the environment through the efficient, economical, and effective acquisition and dissemination of environmental observation data and information. Resources must be applied in a more efficient and effective manner to reduce duplication in today's observing systems, improve coverage, and provide networks to disseminate information and knowledge where and when it is needed around the world.

Over the last few years NOAA has developed enterprise-wide common processes, tools, interfaces, terminology, and standards. NOAA's established enterprise procedures include an approach to prioritizing its verified (and increasingly well validated) requirements as the common metric for evaluating the components of its strategic portfolio across its mission goals. In 2005, NOAA evaluated the complete set of observing systems currently used and proposed by its programs with respect to its prioritized mission-critical requirements, subject to a range of budget constraints. The resulting set of optimal portfolios provides a valuable reference portfolio or target architecture that has supported strategic investment decisions. NOAA is currently in the process of updating this portfolio model for the 2008 programming cycle. The condition, readiness, and vulnerabilities of NOAA's infrastructure have direct consequences on human welfare, economic well being, and the advancement of the understanding of our environment and its management. NOAA's enterprise-wide approach includes integrated management and functional alignment of its high performance computing enterprise along with efforts to develop Agency-wide software infrastructure and community-wide modeling infrastructure capabilities.

IV. Tools

NOAA's mission of monitoring the marine environment requires an extensive global presence. Supporting this mission requires expanded access to ship and aviation platform hours to service buoys, map the coasts, and monitor fisheries. The NOAA database of requirements and capabilities resides in a system: CasaNOSA is the "Home of the NOAA Observing System Architecture". CasaNOSA is an open-source web-based suite of collaborative tools; the system is hosted by the NOAA/NESDIS Office of the Chief Information Officer (CIO) and uses the LAMP Model: (Linux operating system, Apache server, MySQL database, and PHP scripting/programming language). The system houses information ranging from relational databases, documents or software releases depending on the nature of the project and user requirements and specifications. The system contains a selection of tools for input, manipulation and analysis that are widely used during data call periods and survey collection such as the NOAA Planning, Programming, Budgeting and Execution process (PPBES), an exercise where all NOAA Programs, Mission Goals, and Line Offices refine Program Operating Plans (POPs). The uniqueness of the system allows the administrators of each project the ability to control user access, create user friendly views, and easily enter and update data under a flexible mechanism to handle data sets. In order to access the system, a user name and password granted upon request or sponsored by NOAA federal employees is needed. The system security is enhanced by strong authentication procedures to access data. Our current database and toolkit includes the following:¹

CasaNOSA (NOAA Observing System Architecture data base)

Data Collection Surveys/Databases

- **NOAA Observing System Architecture (NOSA)**
- **Consolidated Observation Requirements List (CORL)**
- **Information Management Systems (IMS)**
- **Program Operating Plan (POP)**

Analysis Tools

- **CasaNOSA Analysis System (CAS)**
- **CAS Requirements Tool (CASRT)**
- **Expert Choice - Pair-wise comparison**
- **Portfolio Analysis Machine (PALMATM)**

¹ PALMA is a trademark of The MITRE Corporation, 2008

The Consolidated Observation Requirements List (CORL) is a collection of all of NOAA's environmental observation requirements needed to address the full range of NOAA missions. The CORL provides a comprehensive, standardization media for NOAA Programs to document their environmental observation requirements providing associated data on their requirement's priority level, threshold and objective level needs for attributes such as: geographic coverage, spatial and temporal resolutions and measurement accuracy, and mappings to higher level program outcomes and performance measures. The CORL uses the well established NASA Global Change Master Directory (GCMD) science keyword structure to the fullest extent possible to characterize each environmental requirement at topic, term and variable levels. The same taxonomy is also applied to the NOAA Observing System Architecture (NOSA) and Information Management System (IMS) databases, so that requirements can be matched with capabilities and vice versa.

The NOSA database is a collection of information about observing systems used or proposed by NOAA programs to achieve their outcomes and serve their customers. NOSA includes NOAA's currently funded observing systems, external system or data sources, as well as proposed new systems or enhancements or expansion of current systems. The NOSA database captures the capabilities and costs of current, planned and conceptual observing systems used, or with potential, to address NOAA environmental requirements.

The IMS database captures similar information about NOAA's data management systems. Cross-standardization between the CORL, NOSA, and IMS databases allows fulfillment assessments of NOAA's environmental requirements, identifying observational gaps and investment needs. Annual reviews and updates of all three databases are conducted in the first quarter of each calendar year to assure information integrity and allow for observation requirement analysis assessments to be completed and provided in time for investment recommendations for annual NOAA PPBES cycles.

NOAA has organized itself into the four mission goals and a total of 45 functional areas/programs. On an annual basis, each program creates a new Program Operating Plan (POP) that details what is needed to get the functional job done. The POPs convey the following essential information: resources allocated for the current program, resources needed to meet 100% if the program's mission requirements, products and services those resources will provide, performance measures used by the program to express progress-to-plan and issues, opportunities and alternatives that could affect the program performance and direction.

The CAS is a desktop business intelligence tool that was developed in-house. It is an open-source product (no intellectual property costs) intended for more intensive analysis of data, and requires strong authentication to access any data. CAS retrieves data via web services from multiple sources including CasaNOSA.

The CASRT is Java application that leverages the use of open standards and open source software to deliver powerful and uniformed gap analysis of NOAA's requirements for current and planned capabilities. This tool provides secure access between users and the CasaNOSA server over HTTPS using the Apache Web server. Once the user has been authenticated, CASRT sends two requests, one for the CORL dataset and one for the NOSA dataset. The server then builds a response and sends it back to the user's desktop where Java will parse the data and apply the Multi-attribute Utility Scale Algorithm.

CASRT and the rationale for Multi-attribute Utility Scale (MAUS) Approach

NOAA's operational and science-oriented programs have more than 1000 mission-critical observing requirements and these programs operate and fund over 90 different land, sea, air, and space-based observing systems in addition to obtaining data from a wide range of public, private and International sources. NOAA's observing requirements and capabilities are further delineated in terms of key attributes. The Multi-attribute Utility Scale approach maps performance of observing systems with respect to observing requirements on an attribute-by-attribute basis. These as attributes are: horizontal resolution, vertical resolution, geographic coverage, update frequency, and measurement accuracy. The utility curve used for this mapping could vary from attribute to attribute and requirement to requirement. Different requirement attributes have inherent scales that vary by orders of magnitude; these attributes need to be converted to a common scale with a defined "full satisfaction point." The full satisfaction point is set equal to 90 in this case on a 0 – 100 scale. MAUS uses utility theory to translate observing systems performance at the attribute level to a requirements satisfaction scale that is intuitive and consistent across attributes, requirements, and observing capabilities. However, given the need to simplify and facilitate the initial implementation, the number of different mappings has been limited to four — different requirements or attributes may have different utility curve mappings, users can select utility curves that best reflect the relationship between performance and value.

MAUS Algorithm

Each observing requirement has 3 or 4 primary numerical attributes (R_i) (vertical res, horizontal res, measurement accuracy and update frequency) and 1 categorical attribute (geographic coverage). Each observing system that may contribute to meeting a requirement has corresponding performance or capabilities (C_i). A MAUS utility score (UT) is calculated for each numerical C_i/R_i pair:

$$UT(C_i/R_i) = UT(1) + (1 - (C_i/R_i)N) * (100 - UT(1)) \quad (1)$$

For the geo-coverage attribute a lookup table is used to assign utility scores to Ci/Ri pairs according to the Ri value. The UT for multiple attributes are averaged to determine the overall multi-attribute score, in addition, NOAA programs can set weight factors for each attribute for each requirement. For example, a program may value horizontal resolution more than measurement accuracy for a particular requirement. In these cases, the UT is a weighted average.

Mapping categorical attributes:

Geographic coverage as specified in the NOAA requirements process is inherently categorical (e.g. Global, Hemispheric, or CONUS), not numeric—so a categorical mapping of capability attributes to utility values is needed. Moreover, the value of the geographic coverage of a given capability (such as the hemispheric coverage of a geo-stationary satellite) will vary depending on the geo-graphic coverage category specified in different requirements. For example, if the requirement is for global coverage then a hemispheric capability (say a geostationary satellite) may only provide a value of 50 out of 100. However, for requirements with coverage specified as hemispheric or less, then the satellite would be evaluated as fully satisfying these attributes.

Mapping numerical attributes:

As described above, the remaining attributes (vertical res, horizontal res, measurement accuracy, and update frequency) are numerical and are mapped to a 0-100 utility scale using a mathematical algorithm defined by two parameters; (1) the number that represents “full satisfaction” of the requirement attributes and (2) a “power” parameter that generates the “shape” of the curve (convex or concave; and the steepness of the curve). The description above makes a case for setting the full-satisfaction utility at a value less than 100 (e.g. 90) to allow room on the scale to give some credit or value to capabilities that exceed the full satisfaction level (albeit on a diminishing value curve).

To further simplify the approach, a “linear” utility curve (power parameter = 1) was chosen as the most “neutral” curve (as opposed to either diminishing or increasing value of performance) and has been employed in the demonstration for the three “one-dimensional” numeric attributes: vertical resolution, measurement accuracy, and update frequency. Horizontal resolution, on the other hand is a 2-dimension attribute; increasing the horizontal resolution of an observing system by a factor of 2 (e.g. going from a resolution of 4 km to a resolution of 2 km) increases the amount of data by a factor of four. As a result, the “natural” value curve that linearly reflects the increase in data with decreasing resolution values has a power factor of 2.

V. ANALYSIS OF NOAA’S OBSERVING PORTFOLIO

NOAA generates its prioritization of the relative importance of its mission-critical requirements by means of a value tree model using PALMA™. PALMA is a decision analysis and support tool developed by The MITRE Corporation with federal funding for use by government agencies. PALMA™ is also used to integrate the observing system performance metrics computed in CAS-RT, and investment costs in order to identify a range of optimal observing portfolios over a range of budget constraints. NOAA last compiled an enterprise-wide PALMA model in 2005 and is currently in the process of updating requirement weights and cost information. The 2005 model was assembled without the benefit of CAS-RT, so the 2008 model is expected to provide a more accurate and robust assessment of observing system performance or utility. Data updates and test runs will be completed by the time this paper is presented in September 2008.

The PALMA-based value tree generates requirement priorities or “importance weights” by combining assessments by NOAA’s senior leadership as to the relative importance of its mission goals and programs, and assessments by its program managers and subject matter experts as to the relative importance of each program’s observing requirements. The end result is a decimal value representing the relative importance of each requirement to NOAA. In 2005, NOAA conducted an extensive series of pair-wise-comparison sessions at the goal team and program level to generate these weights and this structure is being substantially updated for 2008. Fig. 1 shows a partial view of the NOAA value tree.

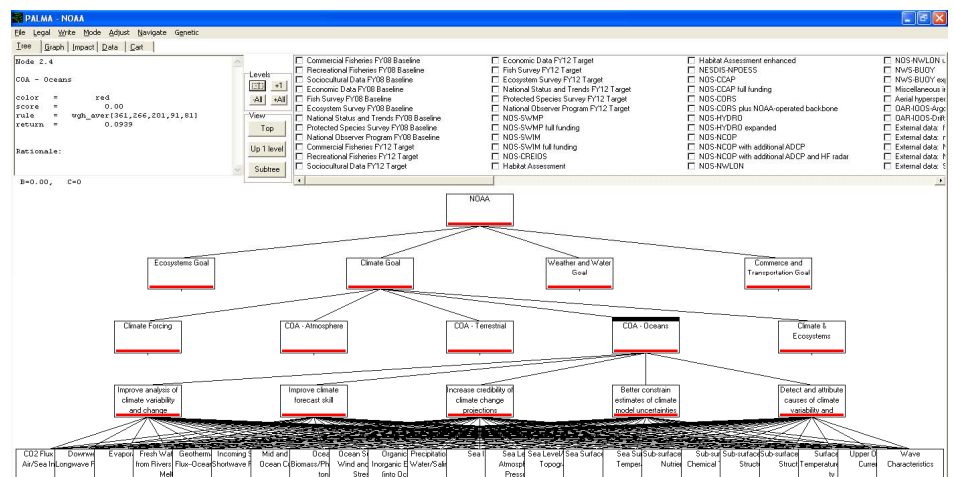


Fig. 1 Screen shot of the PALMA tool showing the value tree with weighted priorities

PALMA™ generates a visual depiction of the optimal collection of investment portfolios at up to 1000 discrete cost constraints. The resulting curve is called the Efficient Frontier. Each point on the frontier represents a portfolio composed of many different observing systems that represents the set of systems that maximizes satisfaction of the prioritized requirements at that cost constraint. An example is shown in Fig. 2. . Generally speaking, we are only interested in the "black dots" that make up the set of efficient portfolios that lie on the frontier itself. The gray dots indicate that there are a large number of possible portfolios--collections of observing systems - - that are in-efficient or sub-optimal. PALMA also provides powerful capabilities to perform "what-if" experiments, such as "forcing in" or "forcing out" investment options, and in asserting dependencies and exclusion rules. These additional constraints usually have the effect of shifting the efficient frontier down or to the right. The gray dots then provide visual clues that indicate that adding constraints results in a new EF that is suboptimal relative to an unconstrained EF. Like all optimization model it has limitations. For example, it does not explicitly address certain technical realities involving the time and resources required to implement certain capabilities; systems cannot be turned on or off at will. However, it does generate a set of target observing portfolios that

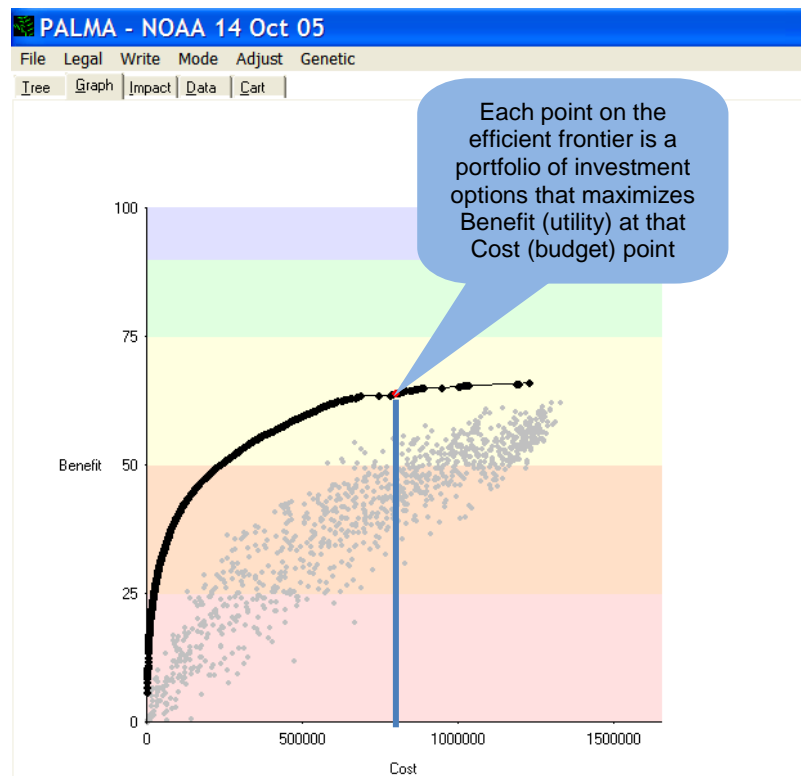


Fig 2. Screen Shot of the Efficient Frontier

incorporates and synthesizes the information contained in the CORL and NOSA databases, incorporates the value judgments of NOAA leadership and subject matter experts. As a result, it provides very valuable guidance on investment strategies to increase overall satisfaction of NOAA's prioritized observing requirements subject to realistic budget and technical constraints. [3]

VI. SUMMARY

Closing the observing system gaps for mission essential needs includes the evaluation of all options for data collection as well as an integrated approach in order for NOAA to implement affordable solutions. Investing effectively in key infrastructure enables NOAA to conduct and fulfill its mission by improving weather warnings and forecasts, enhancing climate monitoring and research, advancing the understanding and management of oceans, coasts and their watersheds, and charting safer, expedient transportation. NOAA's satellites, fleet, aircraft, and other observing systems all face similar problems: expanding mission requirements and finite resources. NOAA currently relies on federal, private sector, university, and international partnerships to help meet requirements for environmental monitoring and data collection in these areas. Even with these valuable partnerships, critical needs in atmospheric, oceanic, climate, solar, and ecosystem observations have gone unfulfilled for many years. An integrated, holistic approach, with the goal of developing observing system architecture, is required to prioritize and appropriately address these limitations while ensuring system continuity for existing products and services. Integration of the vast data and rapidly growing data portfolio poses significant challenges to NOAA's ability to meet its mission. The process of integrating disparate data into an interoperable, multi-discipline knowledge database can be achieved through the development of analysis and intelligent tools.

ACKNOWLEDGEMENT

Thanks to Robert L. Mairs and Robert Reining for their help in editing this paper.

REFERENCES

- [1] Strategic Direction for NOAA's Integrated Global Environmental Observation and Data Management System, pp. 3-9, July 2004.
- [2] Strategic Direction for NOAA's Integrated Global Environmental Observation and Data Management System, p. 15 July 2004.
- [3] Richard A. Moynihan, Robert C. Reining, Patricia P. Salamone, Brian K. Schmidt, "Enterprise Scale Portfolio Analysis at the National Oceanic and Atmospheric Administration (NOAA), accepted for publication in the journal *Systems Engineering* © 2008 Wiley